The Galactic Hitchhiker's Blues

W266 Fall 2016
Final Project Writeup
James King

Background

Everyone knows that galactic travel is hard. Most people don't have to do much of it, but those who travel for a living--salespeople, comedians, bards, hitchhikers--seem stuck with it.

Fortunately, a little known-planet (that was until recently going to be removed for an intergalactic highway) called Earth has a unique and special method for dealing with this kind of inevitable misery. It's a style of music called the Blues.

Motivation

Since there are essentially zero blues musicians outside of Earth, the only reasonable way to make this wonderfully cathartic art available to the innumerable sentient species of the galaxy is to program a machine to establish a set of galactic "blues standards" from which to learn the new genre. Obviously, though, the trivial lamentations of Earthlings will not be appealing to most other creatures. To address this problem, we have generated a family of statistical models of Blues music lyrics from Earth which can be "primed" with any text. When primed with important phrases from galactic literature, lyrics which are more appropriate to space-faring species are generated.

Data Sets

Data set 1: Blues lyrics corpus from the Oxford Text Archive. ~1.4 MB of transcribed blues lyrics in txt format with a few notations indicating song structure. Full data set available here. The song lyrics within corpus have the following characteristics:

Blues Corpus Characteristics

Number of Words	233,761
Number of Tokens	6,722
Number of Unique Bigrams	66,779
Number of Unique Trigrams	150,514
Number of Characters	1,227,760

Data set 2: A list of noteworthy phrases from galactic literature with which to prime the model. (See Appendix B)

Algorithms

Two different approaches to blues song generation were explored

1) Trigrams over the song *words*
2) Long Short-Term Memory (LSTM) neural net trained on the *letters* of the song

Additionally, some special attention was given to automatically formatting the songs into a blues-stanza form with an appropriate number of syllables per line and rhymes in appropriate places.

Trigram Model

Using the Keras abstraction library on top of Tensor Flow a trigram model was generated on the entire blues archive was created. This model was built into a python class which could be called to generate a word based on two words of "context." To get it to generate actual lyrics, some enforcement of Blues conventions was required (described below).

Some of the verses generated with this method were shockingly poetic, while others had a rather entertaining jabberwocky vibe:

```
no back to the station
           in a dive house dance blues
i'm going to get to
           highway forty nine lose
```

```
do better so that mountain
           and i'll drop my backbone
trying to that love my
           little girl drinks shoe telephone
```

LSTM Character Generation

At less than 7000 tokens, the blues corpus is a bit small to train LSTM deep nets at the word level, however at over 1 million characters it's possible to train an LSTM model on the *letters* of the songs in the blues corpus. The best model discovered for blues song generation was a relatively small, single-layer net with a memory of 100 characters. This model was trained on a cloud GPU machine optimized to play nicely with the Tensor Flow back end. The model used for generating the lyrics in this paper took about 20 minutes to train and can generate text in near real time. Example raw output:

```
they calling
       and the time i don't know rooked to time
       that's the way
 going to get
       and and i got the right
       i got to love me
       this home to the girl
       and i believe to know and like the blues
       i said poor mind
 i'm going to be the good mean
```

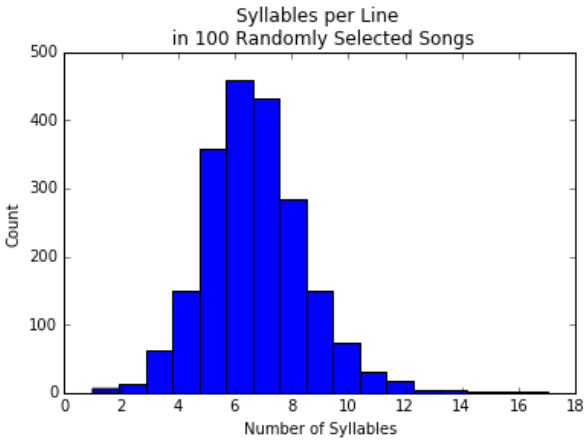
Blues Structure Convention Enforcement

In both the trigram and LSTM cases, the models don't have a sufficient context memory to capture the rhyming schemes and syllable structure at the level of stanzas. Without enforcing this structure, it's very unlikely to appear spontaneously. To adhere to these conventions, a two-layer song formatter was used. These layers were a syllable-per-line processor and a rhyming processor.

Syllable Per Line Processor

Based on data in the blues corpus, the number of syllables per line in blues songs averages about 6.5 with substantive numbers of songs having between 4 and 10 syllables per line.

To structure the words like stanzas, the Carnegie Mellon Pronunciation dictionary is used to count the number of syllables per word. When forming a stanza line, a random number of syllables is chosen from the distribution shown in FIGURE XX. When the number of syllables exceeds this value, a new line is formed. This is repeated for each of four lines and then passed to the rhyming processor.



Rhyming Processor

Neither the Trigram nor LSTM model have sufficient "memory" to capture a rhyming scheme across multiple lines of a song and the corpus is not large enough to train a model a full stanza of memory. Using the Carnegie Mellon University Pronouncing Dictionary (CMUPD) it is very straight-forward to simply substitute a rhyming word for the generated word in a stanza.

This offers a few problems, however. Firstly, the rhyming word won't necessarily even be in the vocabulary of early 20th century musicians. For example, "her" might be paired with the rhyming word "computer." Because rhyme substitution is something of a "hack," it seemed worthwhile to try to make it as ingenuous to the original music as possible. To this end, a 5-level rhyming scheme was developed which tries to stay within the "mood" of the blues corpus wherever possible. The following 5 schemes are traversed in order and the model keeps the first rhyme that works. For example, if there aren't any rhyming words of type 1 or type 2, but the model finds a set of type 3 rhymes, one of the type 3 rhymes is chosen randomly to be inserted.

1. Type 1: Dumb Luck Rhyme: The model happened to have generated a rhyming word--keep it.
2. Type 2: Corpus Pairs Rhyme: Build a bigram model of rhyming pairs used in the corpus. Choose a match from this model.
3. Type 3: Corpus Context Rhyme: Find a rhyming word within the Trigram model with the same context.
4. Type 4: In-Corpus Rhyme: Find any word in the corpus which rhymes
5. Type 5: Repeat Rhyme: Every word rhymes with itself--repeat the word

A close reading of the list above will reveal an apparently missing level between #4 & #5. An obvious approach would be to choose any word in the CMUPD which rhymes. In practice, this tended to include words in the songs which don't "fit," such as the rhyme match of "her" and "computer." Despite the sometimes comical results, the approach was removed.

Take-Aways

There a few take-aways from this project that have been well-seated in my mind. First, neural nets are capable of producing poetry and comedy and nonsense in roughly equal proportions. It makes me ponder about the nature of poetry and comedy. Secondly: rhyming is hard. Fortunately there was some pre-existing infrastructure I could use, otherwise I'd have been in bad shape. Thirdly, it doesn't really matter what you prime the model with. It still generates stuff like the corpus. And finally, letter-based models make interesting words, while word-based models make interesting phrases.

References:

Weide, Robert L. 1998. "The CMU Pronouncing Dictionary." *URL: Http://www. Speech. Cs. Cmu. Edu/cgibin/cmudict*.

Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, et al. 2016. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems." *arXiv [cs.DC]*. arXiv. http://arxiv.org/abs/1603.04467.

Proud, Judith K. 1989. *The Oxford Text Archive*. Vol. 5985. British Library, Research and Development Department.

Chollet, François. 2015. "Keras: Deep Learning Library for Theano and Tensorflow."

Bird, Steven. 2006. "NLTK: The Natural Language Toolkit." In *Proceedings of the COLING/ACL on Interactive Presentation Sessions*, 69–72. COLING-ACL '06. Stroudsburg, PA, USA: Association for Computational Linguistics.

O'neal, Jeff. 2012. "The 42 Best Lines from Douglas Adams' The Hitchhiker's Guide to the Galaxy Series." *BOOK RIOT*. May 25. http://bookriot.com/2012/05/25/the-42-best-lines-from-douglas-adams-the-hitchhikers-guide-to-the-galaxy-series/.

Appendix A: The Song

Another Highway 49 Blues

Lyrics by Recurrent Neural Net #7

Arranged by Penguin Rodeo

My grand-daddy was born
in nineteen hundred and cold
He died in that coal-mine
where he'd been sold

So I'll never go there
And I'll never stay
I'm joining me a jazz band
And I'm going away

(Chorus)
I'm heading off to the station
in my dive house dance shoes
I've got my clarinet and the
highway forty nine blues

[Solo]

You can go holler the vine
My blues are nashville bound
I got both red aces
And some whiskey I found

Someday I'll sit
and ponder the world
i think that i may be poor
but my bullet's my word

(Chorus/Solo/outro)

Appendix B: Song #42 in the Oxford Catalog

Stop, Look, and Listen (recorded 1935)

By Kokomo Arnold

oh stop and listen
hear those bells a-tone
i found my faro
lying on a cooling board

says today has been
a long old lonesome day
seems like tomorrow
mama going to be the same old way

now don't your house look lonesome
when a hearse roll in front of your door
i found my faro
lying on a cooling board

says and it smokes like lightning
and it faro shine like gold
i wouldn't have seen her
not to save nobody's soul

lord then i ain't going down
that big road by myself
if i don't carry you
mama i'm going to carry somebody else

and i followed my faro
to the new burying ground
watch the pallbearers
when they lay my faro down

Appendix C: Selected Priming Phrases

“Funny,” he intoned funereally, “how just when you think life can’t possibly get any worse it suddenly does.”

A common mistake that people make when trying to design something completely foolproof is to underestimate the ingenuity of complete fools.

The reason why it was published in the form of a micro sub meson electronic component is that if it were printed in normal book form, an interstellar hitchhiker would require several inconveniently large buildings to carry it around in.

The last ever dolphin message was misinterpreted as a surprisingly sophisticated attempt to do a double-backwards-somersault through a hoop whilst whistling the ‘Star Spangled Banner’, but in fact the message was this: So long and thanks for all the fish.

“Forty-two,” said Deep Thought, with infinite majesty and calm.

It is a well known fact that those people who most want to rule people are, ipso facto, those least suited to do it. To summarize the summary: anyone who is capable of getting themselves made President should on no account be allowed to do the job.

“He was staring at the instruments with the air of one who is trying to convert Fahrenheit to centigrade in his head while his house is burning down.”

There is an art, it says, or rather, a knack to flying. The knack lies in learning how to throw yourself at the ground and miss.

He hoped and prayed that there wasn’t an afterlife. Then he realized there was a contradiction involved here and merely hoped that there wasn’t an afterlife.

“Nothing travels faster than the speed of light with the possible exception of bad news, which obeys its own special laws.”

In the beginning the Universe was created. This has made a lot of people very angry and been widely regarded as a bad move.

Don’t Panic.